CrossMark

A Catalog of ^{13}CO Clumps from the MWISP in $l = 10°–20°$

Xiaoyu Luo[1,2] , Sheng Zheng[1,3], Zhibo Jiang[4] , Zhiwei Chen[4] , Yao Huang[1,3], Shuguang Zeng[1,3], Xiangyun Zeng[1,3], Rui Zhang[1,3], Chen Long[1,3], Guangrong Zhou[1,3], and Jinbo Hu[1]

[1] Center for Astronomy and Space Sciences, China Three Gorges University, Yichang 443000, China; zsh@ctgu.edu.cn
[2] College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443000, China
[3] College of Science, China Three Gorges University, Yichang 443000, China
[4] Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, China; zbjiang@pmo.ac.cn

Abstract

In this study, we present a catalog of molecular clumps extracted from ^{13}CO ($J = 1 - 0$) emission data of the Milky Way Imaging Scroll Painting (MWISP) project. The data covers the inner Milky Way within the longitude range $10° \leqslant l \leqslant 20°$ and the latitude strip of $|b| \leqslant 5°.25$. The workflow for the extraction of clumps, namely Facet-SS-3D-Clump, consists of two parts: the identification of clump candidates and their verification. First, Facet-SS-3D-Clump employs FacetClumps to identify clump candidates. Subsequently, high-confidence clumps are obtained by cross-matching with the clumps detected by other algorithms, such as dendrogram. Second, these high-confidence clumps are used as prior knowledge to train a semi-supervised deep clustering approach, SS-3D-Clump, which is applied to verify clump candidates detected by FacetClumps, providing confidence levels for the molecular clumps. Finally, the catalog comprising 18,757 molecular clumps was obtained using Facet-SS-3D-Clump, and the catalog is 90% complete above 37 K km s^{-1}. We observe a significant deviation of the mean Galactic latitude for clumps within $|b| \leqslant 2°$ from the midplane, with $b = -0°.110$. We found that 82.3% of the dust clumps correspond to ^{13}CO clumps by matching with Herschel infrared dust clumps. In the future, Facet-SS-3D-Clump will be applied to detect ^{13}CO clumps in the entire MWISP data.

Key words: ISM: molecules – methods: data analysis – stars: formation

1. Introduction

It has been accepted that stars are formed in molecular clouds (MCs; e.g., Shu et al. 1987; Krumholz & McKee 2005; Zinnecker & Yorke 2007; Krumholz et al. 2009; Nejad-Asghar 2011; Könyves et al. 2015; Basu 2016), while the molecular clumps that condensed from the ambient cloud represent a critical step in the star formation process (Hacar et al. 2013). The molecular clump formation also seems to play an essential role in determining the mass of the final star since the distribution of masses among starless cores in a cloud mimics the stellar initial mass function (IMF; e.g., Motte et al. 1998; Alves et al. 2007; Nutter & Ward-Thompson 2007; Könyves et al. 2010; Benedettini et al. 2018; Bresnahan et al. 2018; Arzoumanian et al. 2019). Although great progress has been made in determining the form of the IMF (e.g., Chabrier 2003; Nutter & Ward-Thompson 2007; Olmi et al. 2009; Kroupa et al. 2013; Könyves et al. 2015), a detailed explanation of the form and possible environment depends on understanding the nature and evolution of the molecular clumps (Marsh et al. 2016), especially dense molecular clumps. Thus, the clumps of molecular interstellar medium (ISM) surveys play an important role in understanding the location and mode of star formation in the Milky Way. Since the discovery of CO molecules in the 1970s, numerous survey projects have been

carried out successively, such as the Galactic Ring Survey (GRS; Jackson et al. 2006), the Exeter-FCRAO CO Galactic Plane Survey (Mottram & Brunt 2010), the Milky Way Imaging Scroll Painting (MWISP; Jiang & Li 2013), the Three-mm Ultimate Mopra Milky Way Survey (ThrUMMS; Barnes et al. 2015), the CO Heterodyne Inner Milky Way Plane Survey (CHIMPS; Rigby et al. 2016), the Structure, Excitation, and Dynamics of the Inner Galactic Inter-Stellar Medium (SEDIGISM; Schuller et al. 2017) survey, the Forgotten Quadrant Survey (FQS; Benedettini et al. 2017), etc.

All of those surveys, with different sensitivity and spatial resolution, could help us to detect the distribution of the molecular component with different scales (Benedettini et al. 2021): from dense clouds and filamentary structures (André et al. 2014) to pre-stellar clumps and young stellar objects (YSOs). The MWISP (e.g., Zhang et al. 2014; Sun et al. 2015; Du et al. 2017; Sun et al. 2017; Su et al. 2019) project implemented with the Purple Mountain Observatory Delingha (PMODLH) 13.7 m telescope (Zuo et al. 2011) is dedicated to a new large-scale survey of molecular gas in ^{12}CO ($J = 1 - 0$), ^{13}CO ($J = 1 - 0$), and C^{18}O ($J = 1 - 0$), which targets the northern Galactic Plane within $-10°.25 \leqslant l \leqslant 250°.25$ and $-5°.25 \leqslant b \leqslant 5°.25$ and several other regions of interest. The goal is to develop our knowledge of molecular gas's precise

content and distribution in the Milky Way, the physical and chemical processes converting tenuous interstellar gas to dense molecular gas, and the rules governing the formation of stars from MCs.

With the progress of sky survey projects, unsupervised machine learning-based algorithms for detecting molecular clumps have also emerged (e.g., Rosolowsky et al. 2008; Berry 2015; Luo et al. 2022; Jiang et al. 2023). The dendrogram algorithm, developed by Rosolowsky et al. (2008), is well-suited for illustrating changes in the hierarchical structure of isosurfaces within molecular line data cubes as contour levels vary (Rani et al. 2023). It has been widely used in continuum, atomic hydrogen, and molecular line data (Cheng et al. 2018; Takekoshi et al. 2019; Nakanishi et al. 2020; Zhang et al. 2021). Luo et al. (2022) utilized the local density clustering method, inspired by the concept from Alex Rodriguez (2014), to identify clumps as local dense regions embedded within molecular gas with lower average bulk density, as suggested in (Blitz & Stark 1986; Lada 1992; Bergin & Tafalla 2007). Additionally, a Multiple Gaussian Model (MGM) is utilized to simultaneously fit overlapping clumps, allowing for the extraction of clump parameters. Jiang et al. (2023) presented FacetClumps, a Gaussian facet-based model designed for fitting local surfaces. The extremum determination theorem of multivariate functions is applied to determine clump centers. Then, local regions of the centers are clustered to identify clumps by considering connectivity and minimum distance. The experiments (Jiang et al. 2023) show that FacetClumps exhibits excellent recall and precision rates.

Accumulating survey data and developing automatic detection algorithms for molecular clumps have made it possible to conduct a comprehensive census of these clumps. For example, Wu et al. (2012) expanded the horizon of cold astronomy by studying the Planck cold clumps. Marsh et al. (2016) published a catalog of dense cores of the Taurus star-forming region derived from Herschel SPIRE and PACS observations. Benedettini et al. (2020) presented a catalog of MCs extracted from the FQS ^{12}CO $(J = 1 - 0)$ spectral cubes. Typically, researchers will mitigate false positives through visual inspection, such as Rigby et al. (2019) visually inspecting each CHIMPS clump by three independent reviewers and assigning a reliability flag. Rojas et al. (2022) visually inspected and classified the lens candidates into two catalogs using high-resolution imaging and spectroscopy. At the same time, researchers need to devise meticulous strategies to eliminate potential subjective factors introduced during the visual inspection process, which will require a significant amount of work and time. Therefore, an automated verification method as a substitute for manual verification becomes increasingly necessary in the era of extensive data. Luo et al. (2024) propose a semi-supervised deep clustering method for molecular clump verification, namely SS-3D-Clump, which extracts deep futures of clumps and classifies these features

through a clustering algorithm to obtain a pseudo label. Subsequently, it uses these pseudo labels as supervision to update the weights of the entire network. It can leverage unlabeled samples through semi-supervised learning to enhance the generalization ability of SS-3D-Clump and has shown remarkable performance.

In this paper, we present a comprehensive workflow for detecting and verifying ^{13}CO clumps in the MWISP project. The workflow, namely Facet-SS-3D-Clump, first employs FacetClumps to obtain candidates for molecular clumps and then uses SS-3D-Clump for their verification. To validate the reliability of the workflow, we selected a subregion within $10° \leqslant l \leqslant 20°$ & $-5°\!.25 \leqslant b \leqslant 5°\!.25$ from the MWISP as experimental data, which includes some active high-mass star-forming regions (hereafter HSRs), such as M16 (Hill et al. 2012; Tremblin et al. 2014), M17 (Felli et al. 1984; Chen et al. 2021; Yin et al. 2022), W31 (Beuther et al. 2011; Gama et al. 2016; Maity et al. 2022), W33 (Messineo et al. 2015; Tursun et al. 2022), and W39 (Kerton et al. 2013). At the same time, the infrared dust clumps of Herschel (Elia et al. 2017) also cover the HSR, providing us with an opportunity to validate the Facet-SS-3D-Clump using data from different wavelength bands. The dense molecular cores and clumps are defined as compact ($\sim$0.1 and 1 pc, respectively) and dense ($\gtrsim$10^4–10^5 H$_2$ cm^{-3}) structures (e.g., Williams et al. 2000; Zhang et al. 2009; Ohashi et al. 2016; Motte et al. 2018). Here, we refer to the compact regions traced by ^{13}CO as "clumps." Using Facet-SS-3D-Clump, we obtained a catalog of clumps in the HSR.

The structure and organization of the article are as follows: Section 2 introduces the observations of MWISP. Section 3 describes the details of Facet-SS-3D-Clump that contains the molecular clump extraction algorithm, parameter estimation algorithm, completeness experiments, and SS-3D-Clump to verify clump candidates. In Section 4, we present a catalog containing 18,757 ^{13}CO clumps extracted from HSR and provide a brief statistical analysis of clumps in terms of their spatial distribution and the matching results with Herschel dust clumps. The conclusion is provided in Section 5.

2. Data Introduction

The observations of MWISP use a nine-beam superconducting spectroscopic array receiver, working in the sideband separation mode and employing a fast Fourier transform spectrometer (Shan et al. 2012; Su et al. 2019). The lines of three CO isotopologues, including ^{12}CO, ^{13}CO, and C^{18}O, can be observed simultaneously. All observations are taken in position-switch On-The-Fly (OTF; see Sun et al. 2018) mode. The observing strategy and data reduction adopted for the MWISP Survey are described in detail in Su et al. (2019), and the details on the telescope can be found at http://www.radioast.nsdc.cn/mwisp.php. The ^{13}CO line was observed at the upper sideband with a main beamwidth of $\sim$48″, and the
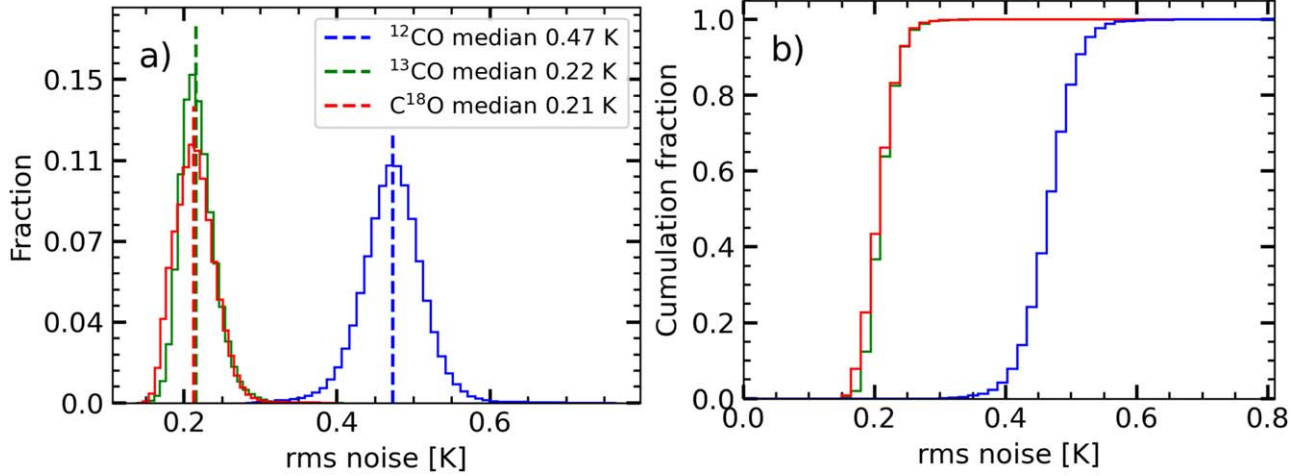
Figure 1. The noise level of the observational data. (a) Histograms of the rms noise levels (σ_{rms}) for ^{12}CO (blue), ^{13}CO (green) and C^{18}O (red). (b) Cumulative distributions of σ_{rms} for three spectra.

^{13}CO and C^{18}O lines were observed at the lower sideband with a main beamwidth of $\sim 50''$. Both bandwidths are 1000 MHz wide with 16,384 channels, resulting in velocity separations of about 0.159 km s^{-1} for ^{12}CO and 0.166 km s^{-1} for ^{13}CO and C^{18}O. The typical noise temperatures including the atmosphere at 110 GHz and 115 GHz are 140 K and 250 K, respectively. A detailed description of the instrument was given by Shan et al. (2012).

The sky coverage of the MWISP project is divided into 10,941 cells for the region. Each cell, being $30' \times 30'$, is scanned along Galactic longitude (l) and Galactic latitude (b) at least twice to reduce the fluctuation of noise. The reasons for choosing a cell size of $30' \times 30'$ are detailed in Su et al. (2019). The sample spacing between adjacent scans is $15''$. Therefore, the MWISP survey fully sampled the mapped area. The OTF raw data were regridded into $30'' \times 30''$ pixels. The data reduction was carried out using the GILDAS software[5] described in detail in Pety (2005). Finally, the three-dimensional (3D) FITS data cubes of each cell were made with a grid spacing of $30''$ for ^{12}CO, ^{13}CO, and C^{18}O ($J = 1 - 0$) lines. The rms distributions are presented in Figure 1. The typical rms noise levels of the spectra are 0.47, 0.22 and 0.21 K for ^{12}CO ($J = 1 - 0$), ^{13}CO ($J = 1 - 0$) and C^{18}O ($J = 1 - 0$), respectively.

The ^{13}CO integral $l - b$ map of HSR is shown in Figure 2, which is integrated from $v_{LSR} = -5$ to 80 km s^{-1}. The integrated intensity threshold is 2 times σ_{rms}, where σ_{rms} is the rms noise per velocity channel. There are several renowned star-forming regions in HSR. Surveying the molecular clumps in HSR can provide a substantial sample for studying different

evolutionary stages of molecular clumps. It is also a crucial step in understanding the early stages of stellar evolution.

## 3. Generation of Clump Catalogs from MWISP

Star-forming regions are messy and chaotic environments, with structures on many scales (Wurster & Rowan 2023). The entire region is typically referred to as a cloud, dense regions embedded within the cloud are clumps, while the very dense regions in the clumps are cores (Bergin & Tafalla 2007). There is general agreement of the three terms, although there is ambiguity among the specific definitions and divisions between the levels. For example, Rathborne et al. (2009) identified MCs and clumps using the ClumpFind algorithm, utilizing ^{13}CO ($J = 1 - 0$) emission line data observed by the GRS (Jackson et al. 2006). Takekoshi et al. (2019) analyzed the statistical properties of C^{18}O ($J = 1 - 0$) clumps in the Cygnus X cluster-forming region using data from the Nobeyama 45 m radio telescope, with the clumps identified through the dendrogram algorithm (Rosolowsky et al. 2008). Liu et al. (2022) refer to centrally concentrated structures as clumps and treat dendrogram-defined leaves accordingly.

The main characteristics of the molecular clumps are local intensity enhancement and different shapes. Meanwhile, they are embedded in lower average bulk density molecular gas (Blitz & Stark 1986; Lada 1992). The clumps (or cores) are empirically defined as the regions with concentrated, enhanced intensity in a data cube. It is worth noting that ^{13}CO can only trace structures of specific density, which falls within the range of 10^3–10^4 cm^{-3}. This limitation results in ^{13}CO only revealing local structures within MCs. Therefore, in this study, we define the structures traced by ^{13}CO data as clumps.
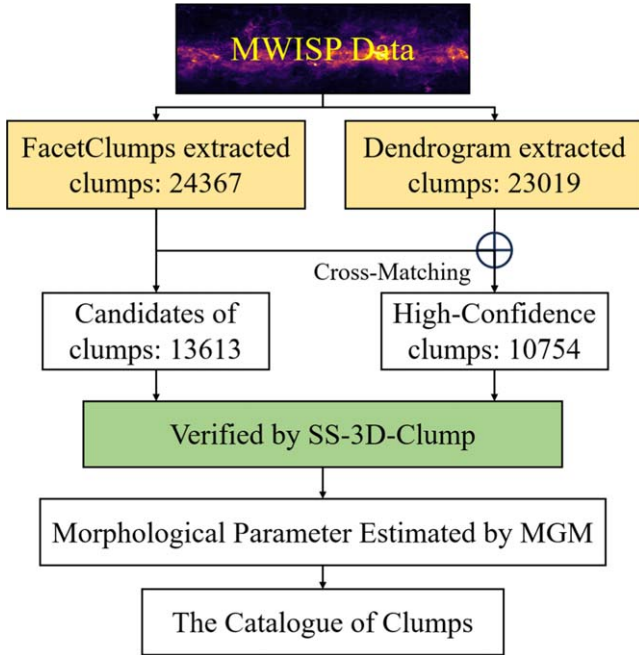
The workflow of Facet-SS-3D-Clump for obtaining the catalog of molecular clumps is shown in Figure 3, in which we

---

5 http://ascl.net/1305.010 or http://www.iram.fr/IRAMFR/GILDAS
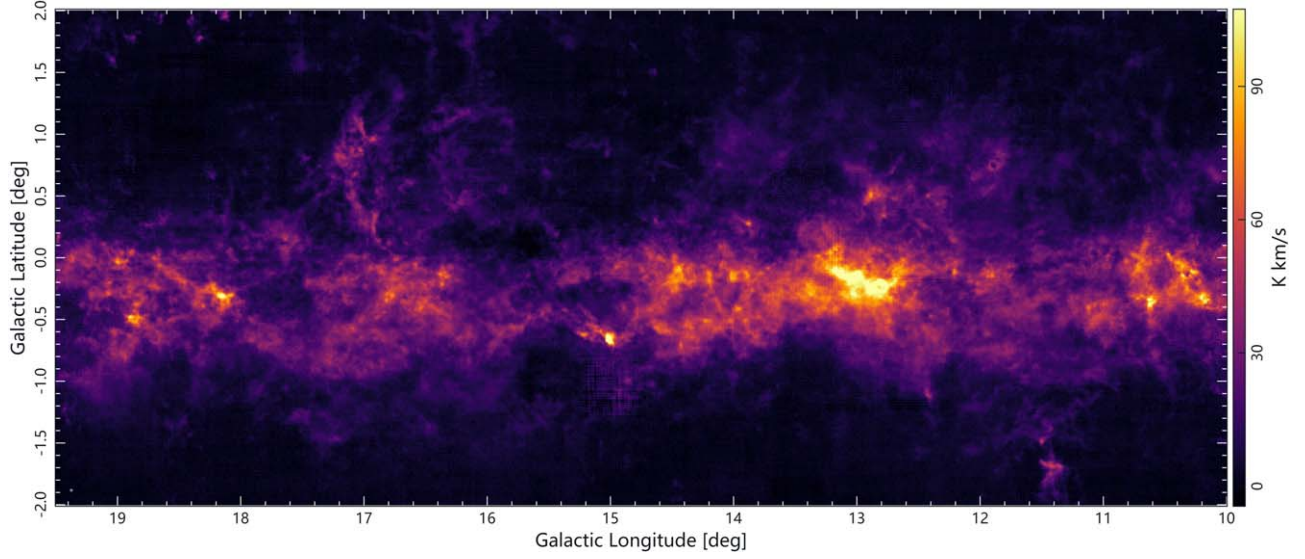
Figure 2. The intensity maps of ^{13}CO emission for the HSR in velocity ranges from -5 to 80 km s^{-1}. The integrated intensity threshold is $2 \times \sigma_{rms}$, where σ_{rms} is the rms noise per velocity channel.

Figure 3. Workflow of obtaining the catalog of clumps.

utilized the dendrogram (Rosolowsky et al. 2008) and FacetClump (Jiang et al. 2023) algorithms to detect ^{13}CO data. Subsequently, we performed cross-matching to obtain clumps with high confidence. Simultaneously, these high-confidence clumps were used as prior knowledge to train an SS-3D-Clump model for verifying the remaining clumps, resulting in confidence levels for molecular clumps. At the same time, we

used an MGM described in Luo et al. (2022) to simultaneously fit overlapping molecular clumps and derive the geometric morphological parameters of those clumps. Finally, we obtained the catalog of molecular clumps, which includes the positions, morphological parameters, and confidence levels associated with the clumps.

To avoid the difficulty of detecting clumps from the entire MWISP data, which is computationally expensive, we perform the clump extraction with a strategy that processes sub-cubes and global stitching (see Figure 4). A sub-cube spans 120 pixels in the longitude and latitude directions, i.e., $\sim 1°$, given the MWISP data pixel size of $30''$. The candidates are rejected from the first catalog since their centers are outside the local region of the sub-cube (see Figure 4). Then, the clumps belonging to the entire region are obtained by retaining the clumps whose centers are in the local regions of all sub-cubes. As shown in Figure 4, the clumps labeled with designations a, b, c are in sub-cube 1, while the center of the clump c marked by a red dot is out of the local region of sub-cube 1. Thus, clump c colored with gray is removed from sub-cube 1 and clumps a and b colored with blue are retained in the catalog of sub-cube 1. In sub-cube 2, the center of clump b is out of the local region, while that of clump c is in the local region. The clump c is retained in the catalog of sub-cube 2.

3.1. Clump Extraction Algorithm

Dendrogram (Rosolowsky et al. 2008) implemented in astrodendro[6] is an abstraction of the changing topology of the isosurfaces as a function of contour level, which uses a tree
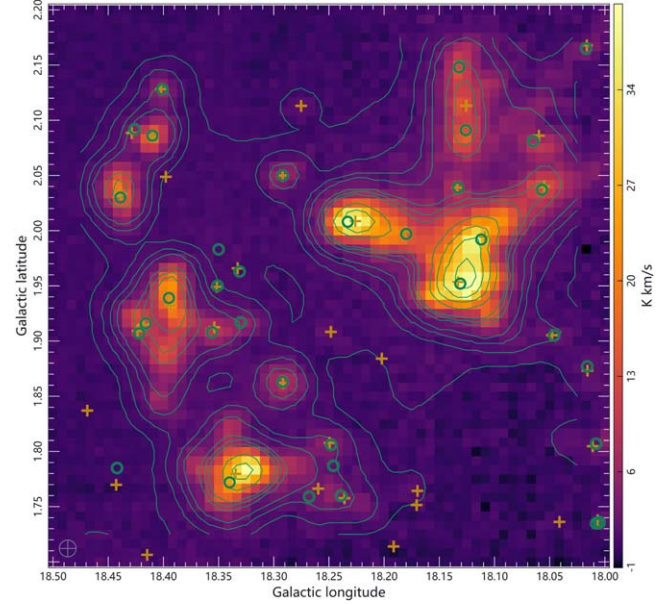
---

[6] https://dendrograms.readthedocs.io/en/stable/
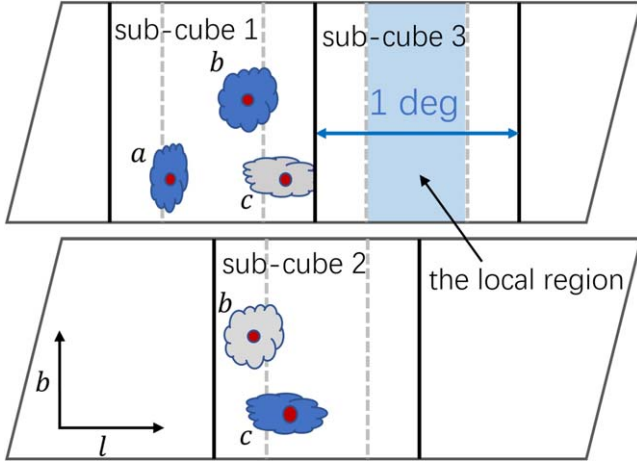
Figure 4. Scheme of processing sub-cubes and global stitching for the entire MWISP data. Sub-cube edges are indicated with black vertical lines, which span all approximately 1° in longitude (latitude is the same). The light blue region surrounded by vertical dotted lines marks the local region of the sub-cube. Clumps *a* and *b* in sub-cube 1 and clump *c* in sub-cube 2 are retained, according to whether the center of the clump falls into the local region.

Figure 5. Dendrogram and FacetClumps were applied to the same ^{13}CO data; the green circles mark the centroids of clumps detected by FacetClumps, while the orange plus symbols represent the results by dendrogram.

diagram to describe hierarchical structures over a range of scales in a two-dimensional (2D) or 3D datacube (Zhang et al. 2021). There are two types of structures returned in the results: leaves, which have no substructure, and branches, which can split into multiple branches or leaves. There are two main parameters in the algorithm: T_{min} and ΔT. T_{min} is the minimum value to be considered in the data set. In the fiducial case, we adopt $T_{min} = 3\sigma_{rms}$, where σ_{rms} is the local noise level. ΔT describes how significant a leaf has to be in order to be considered an independent entity. We adopt a fiducial value of $\Delta T = 2\sigma_{rms}$, which means a clump must have a peak flux reaching $5\sigma_{rms}$ above the noise.

FacetClumps[7] was developed by Jiang et al. (2023), which uses the Gaussian facet model to fit local surfaces and employs the extremum determination theorem of multivariate functions to determine clump centers. Based on the identified clump centers, FacetClumps clusters the regions near the centers by considering connectivity and minimum distance, thereby obtaining and detecting molecular clumps. The FacetClumps algorithm uses default parameters in this paper (see Appendix C in Jiang et al. 2023). By cross-matching the detection results of the two algorithms, "extremely" high-confidence clumps can obtained. As shown in Figure 5, drawn using the Cube Analysis and Rendering Tool for Astronomy (CARTA; Comrie et al. 2021), the green circles and blue plus signs represent the centroids of clumps detected by FacetClumps and dendrogram, respectively. When the deviations of the peak positions of two clumps detected by two algorithms are less than 2 pixels in all

three directions, corresponding to deviations less than 1′ in Galactic longitude and latitude and less than 0.32 km s^{-1} in the velocity axis in the actual data, they are considered to be the same clump. Among them, FacetClumps and dendrogram detected 24,367 and 23,019 clumps in HSR, respectively, with 10,754 clumps matched between the two.

3.2. Verified by SS-3D-Clump

To minimize false positives, both Rigby et al. (2019) and Rojas et al. (2022) employed visual inspection by independent reviewers, assigning reliability flags for CHIMPS clumps and classifying lens candidates into two catalogs using high-resolution imaging and spectroscopy, respectively. Additionally, to minimize the introduction of subjectivity in manual verification, the researcher makes every effort to control this uncertainty during the verification process. For example, Rojas et al. (2022) re-displayed the previously classified ring galaxies during the verification process as a consistency check (users should re-classify them as rings, or at least not classify them as lenses). Therefore, manually verifying candidates requires significant effort and time, especially for large amounts of data.

Therefore, an automated verification method as a substitute for manual verification becomes increasingly necessary in the era of extensive data. Inspired by the success of supervised deep learning in galaxy classification (Zhu et al. 2019; Cheng et al. 2020; Lukic et al. 2019; He et al. 2021; Gupta et al. 2022), we integrate limited labeled data with a deep learning approach. Utilizing semi-supervised deep learning, namely SS-

---

[7] https://github.com/JiangYuTS/FacetClumps
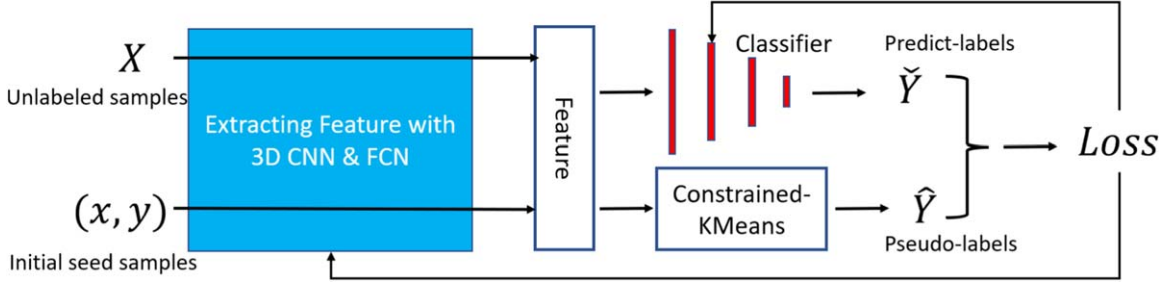
Figure 6. The structure of SS-3D-Clump is illustrated in the figure. The proposed method refers to iteratively clustering deep features and using cluster assignments as pseudo-labels to learn the parameters of the 3D Convolutional Neural Network (CNN) and classifier networks.

3D-Clump (Luo et al. 2024), we verify the clump candidates. As shown in Figure 6, the deep features of the clump candidate are extracted by the feature extraction part of SS-3D-Clump (Luo et al. 2024). Utilizing the deep features as a foundation, the Constrained-KMeans (Basu et al. 2002) algorithm efficiently generates pseudo-labels for candidates using small labeled samples as seeds. The SS-3D-Clump classifier also assigns predict-labels to the candidates based on their deep features. The difference between both labels could be used to optimize the parameters of the SS-3D-Clump model. SS-3D-Clump iteratively groups the deep features with a standard clustering algorithm and uses the subsequent assignments as supervision to update the network weights of SS-3D-Clump.

Figure 7 illustrates the curve of the Normalized Mutual Information (NMI) changing with the epochs during the training process of SS-3D-Clump. NMI is a metric used to assess the similarity of two data sets or the performance of clustering algorithms. The range of NMI typically falls between 0 and 1, where 1 indicates complete similarity between two clustering results, and 0 signifies no resemblance. During the training process, NMI curves can be utilized to observe how the algorithm's performance changes with increasing training epochs, facilitating the optimization of the model training process. Measuring NMI between the clusters at epoch $t-1$ and t provides insights into the actual stability of SS-3D-Clump. From Figure 7, it can be observed that, by the 15th epoch of training, the NMI increases to around 0.97 and remains stable. This indicates that SS-3D-Clump is experiencing fewer reassignments, and the clusters are stabilizing.

Since SS-3D-Clump belongs to semi-supervised deep learning, the verification of molecular clumps is completed simultaneously with the model training. Utilizing the stabilized SS-3D-Clump after training, we conducted verification on the candidates detected by FacetClumps, excluding those with confidence levels below 0.8. In the end, SS-3D-Clump excluded 5792 out of the 24,367 molecular clump candidates detected by FacetClumps, resulting in a catalog containing 18,575 molecular clumps. The histograms of peak intensities of the molecular clumps are illustrated in Figure 8. The green
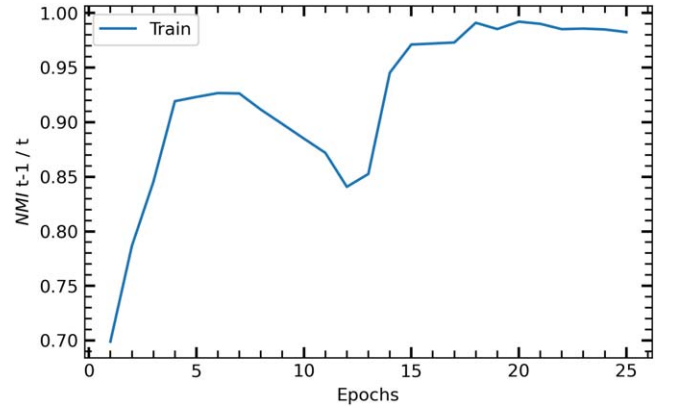
Figure 7. Evolution of cluster reassignments (NMI) in each clustering iteration of SS-3D-Clump.

histogram depicts the peak value distribution of clumps retained by SS-3D-Clump, while the red one shows the cases excluded by SS-3D-Clump. It can be inferred from Figure 8 that the distribution of retained clumps has a sharp rise spike with position of about 1.5 K. After passing the peak position, the distribution drops rapidly and slowly descends. Appendix A presents a subset of the results obtained through SS-3D-Clump verification, which contains integrated maps of molecular clumps on the $l - b$ plane along with their associated confidence levels.

3.3. Morphological Parameter Estimation

For blended source pairs, the Gaussian fitting results in a larger size estimate, whereas in cases where two or multiple sources are resolved by the detection algorithm, simultaneous fitting of multiple components enables better separation (Molinari et al. 2016). Therefore, we use the MGM[8] described in Luo et al. (2022) to simultaneously fit overlapping molecular clumps and derive the geometric morphological parameters of those clumps. The integrated map shown in Figure 9 represents

---

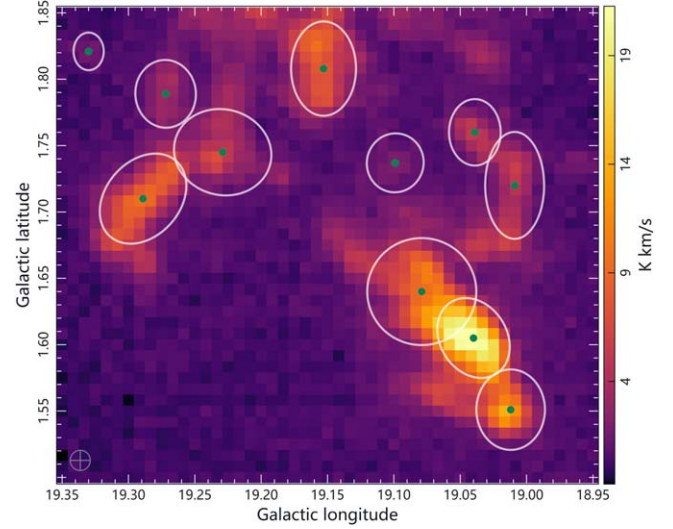[8] https://github.com/Luoxiaoyu828/LDC-MGM
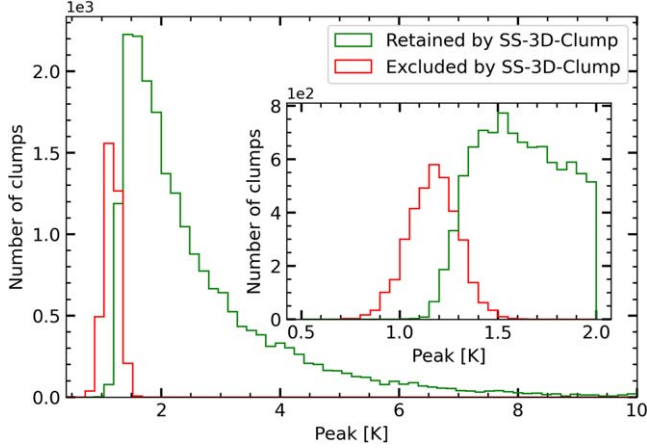
Figure 8. Histogram illustrating the distribution of peak values for molecular clump candidates detected by FacetClumps and verified by SS-3D-Clump. The green bin markings represent the distribution of retained clumps, while the red ones represent those excluded by SS-3D-Clump.

Figure 9. The figure displays an integrated map of a local region in HSR with an integrated velocity range of 25.6–28.1 km s^{-1}. Morphological parameters for the molecular clumps detected by FacetClumps are estimated using MGM. The green dots indicate the centroids of clumps, and the white ellipses are drawn based on the major axis, minor axis, and rotation angle of the clumps.

a local region in HSR with an integrated velocity range of 25.6–28.1 km s^{-1}. For the molecular clumps detected by FacetClumps, morphological parameters are estimated using MGM. The green dots signify the centroids of clumps, and the white ellipses are drawn based on the clumps' major axis, minor axis, and rotation angle.

3.4. Completeness

The limitation of the telescope sensitivity causes a low quality clump to be ignored. Other indicators of the algorithm are the completeness and the detection rate above the limitation. The "completeness limit" here refers to the total flux or mass above which a clump can be detected at a certain level with an algorithm. The smaller and weaker molecular clumps are, the less likely they are to be detected. To quantify the degree of completeness of the extracted clumps, we carried out an extensive set of synthesis data experiments by injecting simulated clumps into the ^{13}CO $(J = 1 - 0)$ data in HSR. Completeness experiments for ^{13}CO were conducted by partitioning the entire HSR into 16 subregions, collectively covering the entire HSR.

Via multiple injections, we injected 36,000 simulated clumps modeled as 3D ellipsoidal Gaussians for each subregion, with their axis size and intensities randomly distributed within specific ranges. The peak intensity value of those simulated clumps takes values from 0.7 to 15, while the size of the clump in the velocity axis takes values from 2 to 5 and the size in the Galactic longitude and latitude axes takes values from 1.5 to 4. In this way, we were able to assess the capability of recovering a statistically comparable population of clumps from the designated region. These simulated clumps were randomly

spread on the map, with the only constraint being to avoid overlap in their positions.

The synthesis data were processed with FacetClumps, adopting the same parameters as used for real data of HSR, and the outputs were compared with the truth table of the simulated clumps. The number of clumps within each total flux interval was counted, and the completeness was obtained. An example of the recovery fraction as a function of the integrated flux density of the simulated clumps in different latitude intervals in HSR is shown in Figure 10. From Figure 10, it can be observed that the completeness curves decrease significantly when the flux of clumps is from 40 to 20 K km s^{-1} within different latitude intervals. When the completeness reaches 90% for the entire region, the total flux of clumps is approximately 37 K km s^{-1}. In Figure 11, the dashed line represents the estimated completeness limit as a function of Galactic latitude. At the same time, the histogram shows the distribution of missed simulated clumps in Galactic latitude during the synthetic data experiment. Around 0°, specifically in $-1°\!.5 \leqslant b \leqslant 0°\!.5$, the dashed line exhibits a noticeable increase, corresponding to a peak in the histogram at this position. It is explained by the brighter emission at lower latitudes, making detecting fainter objects a more challenging task.

The completeness limits reported in Figure 11 should be seen as conservative because they are determined by spreading the synthetic clumps randomly over each entire region. However, the background has non-uniform characteristics in each region, decreasing toward the north and south Galactic directions as the
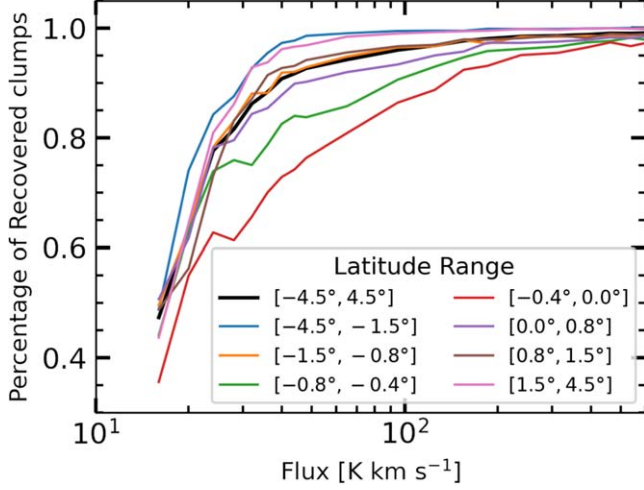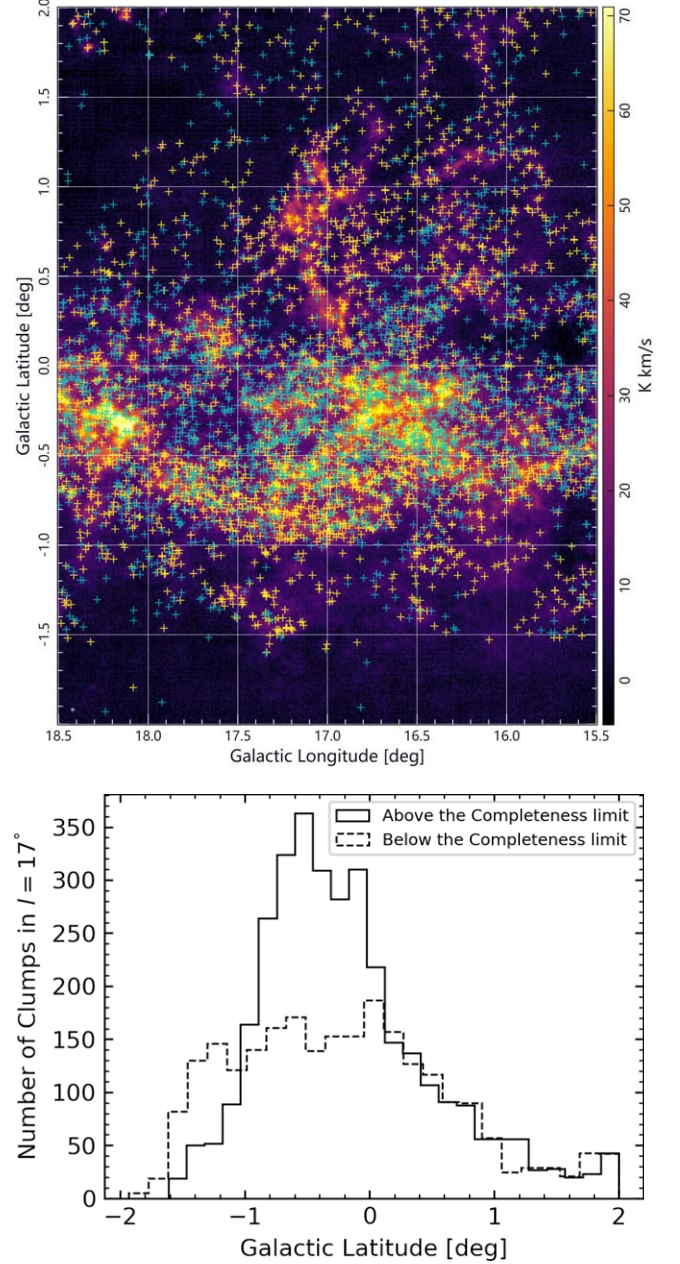
Figure 10. Completeness fractions as a function of flux density in different latitude intervals within HSR, corresponding to molecular clumps in the extracted catalog with statistically the same sizes.
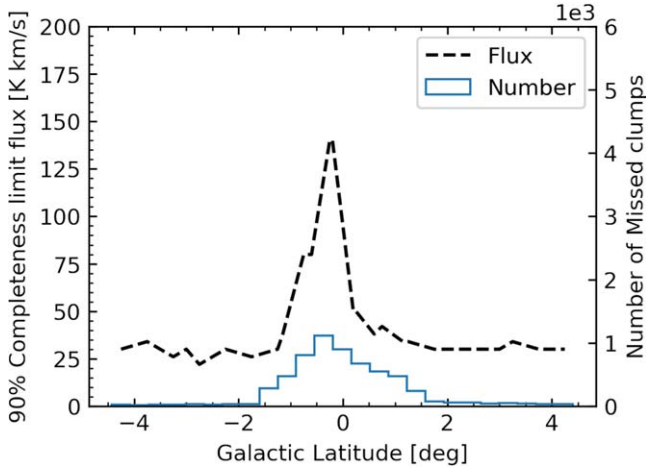
Figure 11. Ninety percent completeness limits in flux density for a population of clumps with the same distribution of sizes as the one extracted by FacetClumps as a function of Galactic latitude. The significant increase in the completeness limit near low Galactic latitudes is due to the brighter background emission in HSR.


Figure 12. The spatial distribution of molecular clumps above and below the completeness limit and their distribution in Galactic latitude. Upper panel: the integrated ^{13}CO map of MWISP at $l = 17°$. Superimposed are the clumps detected by FacetClumps. The yellow crosses indicate the clumps with fluxes above the completeness limit, while the blue crosses indicate the clumps below the completeness limit. Lower panel: Histograms of latitude distributions for clumps, above (solid line) and below (dashed line) the completeness limit.

Galactic latitude increases. In Figure 12, a typical example is shown, with the upper panel showing the integrated ^{13}CO map of MWISP centered at $l = 17°$. Superimposed are the extracted clumps with integrated fluxes above (yellow crosses) and below (blue crosses). The lower panel of Figure 12 further depicts the latitude distribution of the two groups of clumps, also signified with solid and dashed lines for sources above and below the confusion limit respectively. The two groups of clumps exhibit

subtle differences in their spatial distribution, with clumps brighter than the completeness limit mainly concentrated in the Galactic latitude range of $-1°$ to $0°.6$. In contrast, the distribution of the other group of clumps is more uniform.
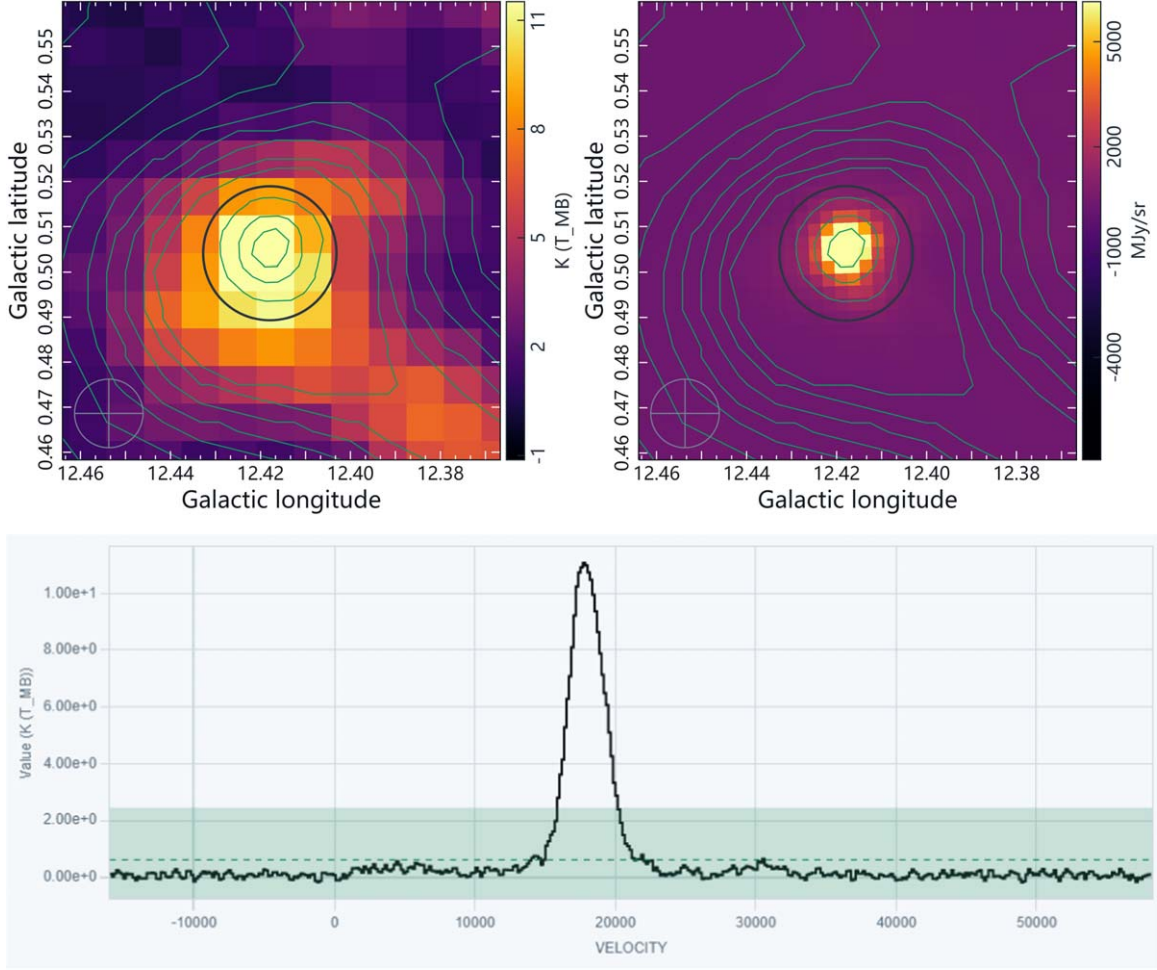
Figure 13. The one-to-one matching results of Herschel 350 μm dust clump and ^{13}CO clump. The upper-left subplot displays the channel map with the maximum intensity of the ^{13}CO clump, with the corresponding Herschel dust clump located in the upper-right subplot. The lower subplot illustrates the average spectrum of the ^{13}CO clump, and the averaging range is delineated by the black circle in the upper-left subplot.

Table 1
Description of Columns in the Catalog

Column Name	Units	Symbol	Explanations
ID	…	…	Designation of the clump (MWISPLLL.lll+B.bbb+VVV.v)
GLON	deg	l	Galactic longitude
GLAT	deg	b	Galactic latitude
Velocity	kms^{-1}	v_{LSR}	the central radial velocity with respect to LSR
FWHMA	arcsec	S_M	FWHM of the clump along axis a of the elliptical Gaussian as determined by fitting engine
FWHMB	arcsec	S_m	FWHM of the clump along axis a of the elliptical Gaussian
FWHMV	km s^{-1}	S_v	FWHM of the clump in the velocity direction
PA	deg	θ	Position angle
Peak	K	T_{peak}	The Peak value of clump
Flux	K km s^{-1}	W	The total integrated flux of clump
S/N	…	r	Signal-to-noise ratio obtained by dividing Peak by the residual rms
Fitting Error	E	K	E is calculated by the root mean square error between the fitting and the actual value
Confidence	…	α	The confidence associated with the clumps verified by SS-3D-Clump
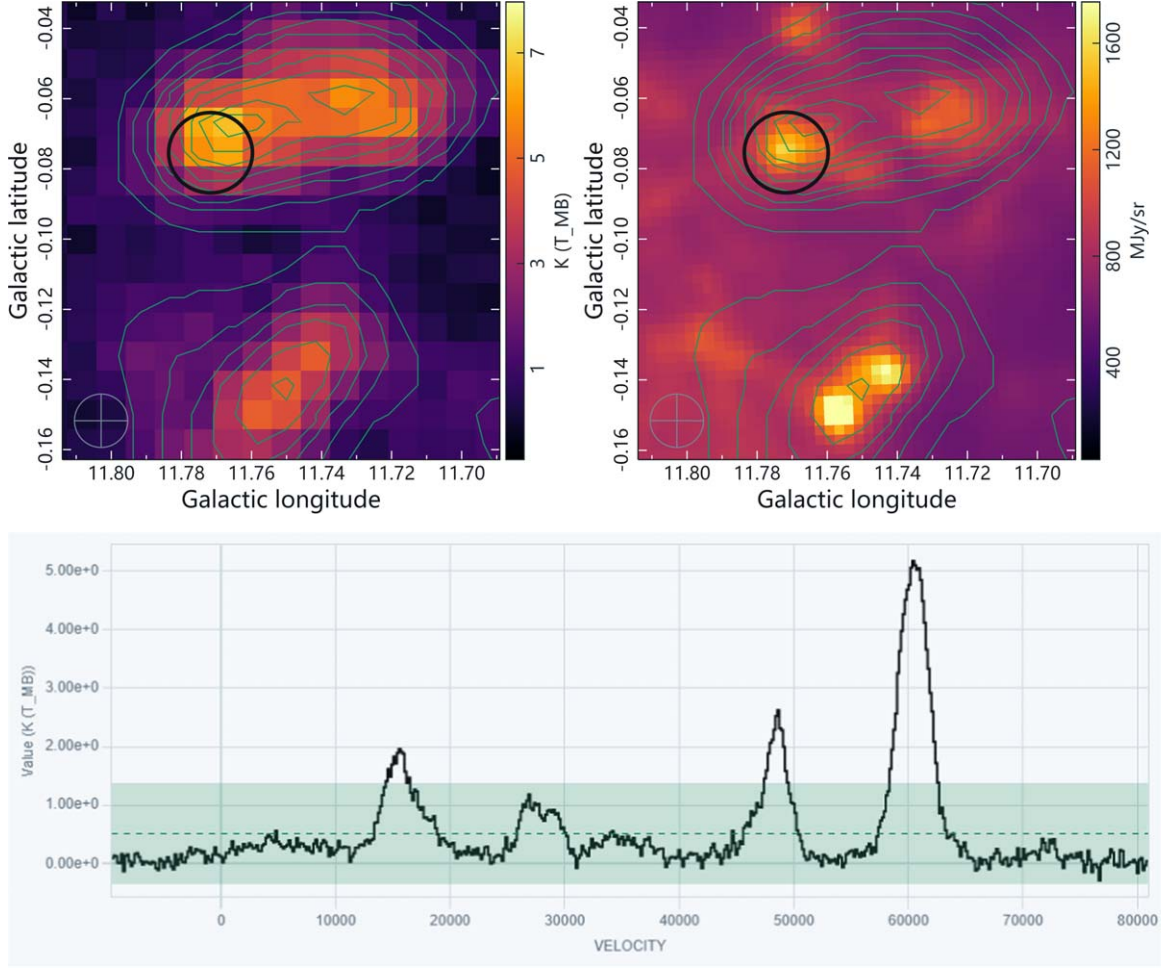
Figure 14. The figure illustrates results similar to Figure 13, with the distinction that it showcases a Herschel dust clump matched to multiple ^{13}CO clumps.

4. Results

4.1. Detailed Contents of Clump Catalog

A total of 18,575 ^{13}CO clumps are extracted in HSR. Table B1 in Appendix B presents a part of the clump catalog, while the full catalog can be obtained online (https://www.scidb.cn/en/s/qEfe2m). In the table, columns (2) and (3) are the positions of the clumps in galactic coordinates, and column (4) is the central radial velocity with respect to the local standard of rest (LSR). Columns (5) and (6) represent the major and minor sizes of the fitted ellipse on the sky plane, respectively. Column (7) represents the FWHM size in the velocity direction. Column (8) is the tilt angle on the sky plane, calculated from the north Galactic pole to the major axis of the clump in the counterclockwise direction. Columns (9) and (10) list the clump's peak intensity and total integrated intensity value, respectively. Column (11) represents the peak signal to noise ratio (S/N) of the clump. Column (12) is the fitting error, which is calculated by the rms error between the fitting value and the actual value.

Column (13) is the confidence levels associated with the clumps, which are verified by SS-3D-Clump. Names, units, symbols, and detailed descriptions of the columns in the molecular clump catalog are provided in Table 1.

4.2. Cross Match with Hi-GAL Sources

Herschel infrared Galactic Plane Survey (Hi-GAL) is a large-scale survey of the Galactic plane (Molinari et al. 2010), performed with Herschel in five infrared continuum bands between 70 and 500 μm using the Herschel Space Observatory (Pilbratt et al. 2010). The first public release of the high-quality products from the Hi-GAL survey was presented by Molinari et al. (2016), in which $\sim 10^5$ compact sources are identified in each band. Elia et al. (2017) merged the sources of the five bands mentioned above to obtain a band-merged catalog containing 100,922 sources with a regular SED, 24,584 of which show a 70 μm counterpart and are considered protostellar.
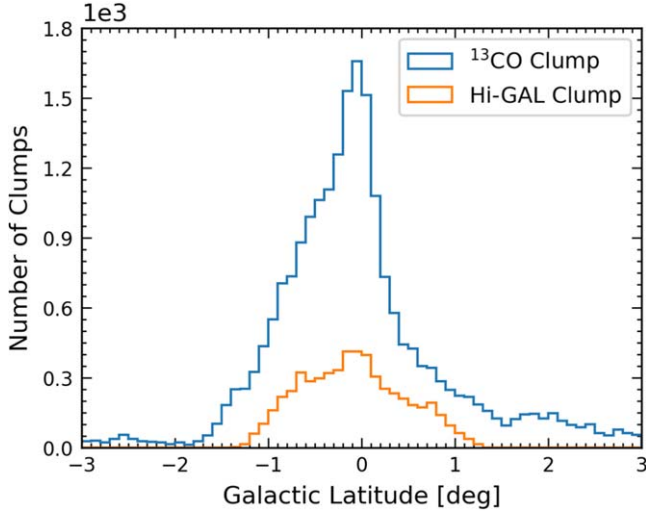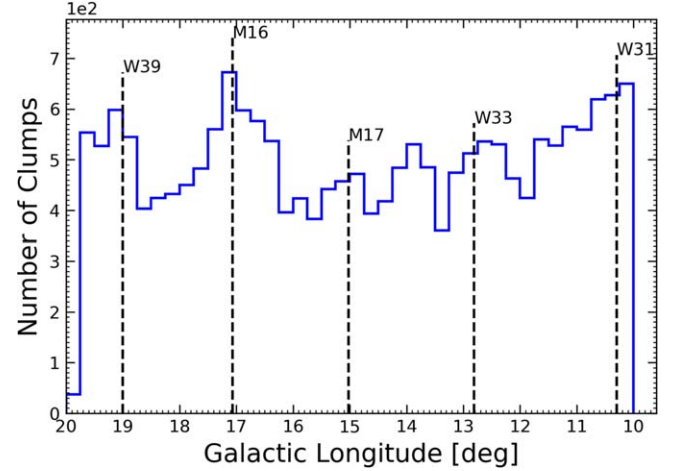
Figure 15. Distribution of molecular clumps along Galactic latitude in HSR. The blue histogram represents the distribution of molecular clumps detected by Facet-SS-3D-Clump, while the orange histogram represents the distribution of sources in the Herschel survey.

Figure 16. The distribution of ^{13}CO clumps along the Galactic longitude, where the dashed line signifies the star-forming region in the HSR. The figure reveals a local clustering of ^{13}CO clumps near the longitude of the star-forming region, indicating the consistency of clumps detected by Facet-SS-3D-Clump.

Within the Galactic longitude range $10° \leqslant l \leqslant 20°$, there are 8517 band-merged sources, and through matching with ^{13}CO clumps, 7154 sources have been identified as corresponding to ^{13}CO clumps. Among them, 2200 sources are one-to-one matches with ^{13}CO clumps, while 4954 sources match multiple ^{13}CO clumps at different velocities. As shown in Figure 13, the figure illustrates the one-to-one matching results of the Herschel 350 μm dust clump and ^{13}CO clump. The upper-left subplot of Figure 13 shows the channel map with the maximum intensity of the ^{13}CO clump and the corresponding Herschel dust clump is in the upper-right subplot. The lower subplot of Figure 13 represents the average spectrum of the ^{13}CO clump, with the averaging range indicated by the black circle in the upper-left subplot. Looking at the blue contours of ^{13}CO in Figure 13 and the overlay of these contours onto the dust clumps, the positions of the two correspond very well. As shown in Figure 14, similar to Figure 13, the only difference is that the displayed results depict a Herschel dust clump matched to multiple ^{13}CO clumps. The figure reveals a similarity between the ^{13}CO data and Herschel dust emission. In the one-to-multi matching results, the average spectrum of ^{13}CO data at the positions of Herschel dust clumps exhibits multiple peaks, which can explain why the number of ^{13}CO clumps is greater than that of Herschel dust clumps. The upper-left subplot of Figure 14 displays the channel map of the maximum intensity in the average spectrum.

4.3. Spatial Distribution

A comparison with Hi-GAL dust clumps (Elia et al. 2017), covering the inner Milky Way in the longitude range

$10° \leqslant l \leqslant 20°$, shows substantial similarities in the source count distributions. The latitude distribution of the HSR clumps and the Hi-GAL compact sources is reported in Figure 15. The blue histogram represents the distribution of molecular clumps obtained by Facet-SS-3D-Clump, while the orange represents the distribution of dust clumps in Hi-GAL (Elia et al. 2017). Both histograms peak at slightly negative values, and the median value for HSR clump latitude is $\sim -0°.118$ below the nominal midplane. Also, the mean Galactic latitude for clumps within $|b| \leqslant 2°$ is significantly below the midplane: $b = -0°.110$. From the histograms, it can also be observed that the distribution of molecular clumps exhibits an asymmetrical pattern. However, we cannot rule out the asymmetric distribution with respect to the central plane in the studied field. As Molinari et al. (2016) pointed out, this may be attributed to an incorrect assumption about the vertical position of the Sun in the Milky Way.

The longitude distribution of ^{13}CO clumps is shown in Figure 16, with the dashed line representing the star-forming region in HSR. From the figure, it can be observed that there is a local clustering of ^{13}CO clumps near the longitude of the star-forming region, which also reflects the consistency of clumps obtained by Facet-SS-3D-Clump. The top panel of Figure 17 shows the positions of clumps in the $l - b$ plane. The bottom of Figure 17 presents the centers of clumps in the $l - v$ plane. The yellow, red, green, and purple lines correspond to the spiral arms of Aql Rift, Sagittarius, Scutum, and Norma, respectively. Dashed lines indicate the nearness of the arms, while solid lines represent their farness.
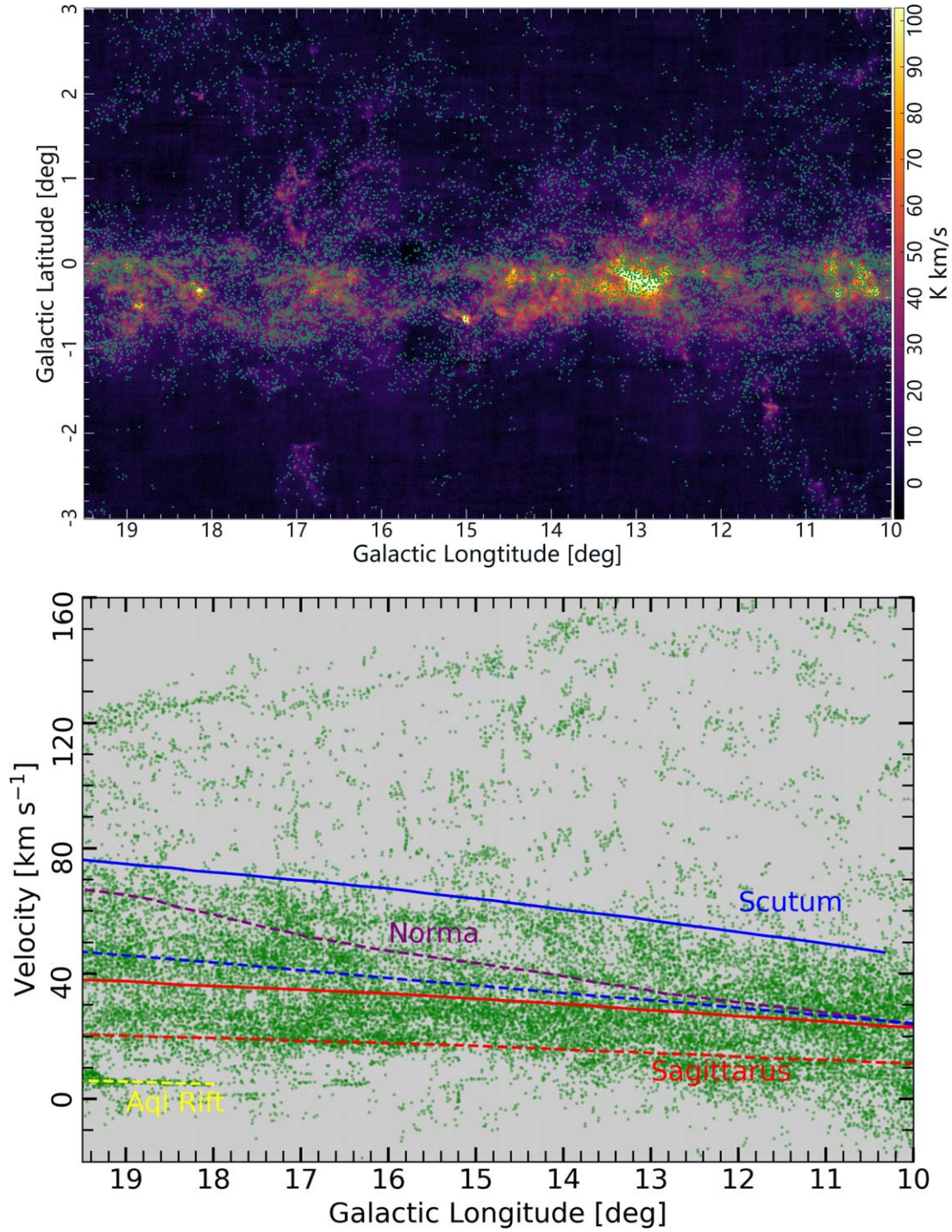
Figure 17. The distribution of clumps in both spatial and velocity dimensions. Top: the positions of clumps in the $l-b$ plane. The green dots indicate centroids of clumps. Bottom: the positions of the clumps are shown on the $l-v$ plane. The yellow, red, green, and purple lines represent the Aql Rift, Sagittarius, Scutum, and Norm's arms, respectively. Dashed lines indicate the nearness of the arms, while solid lines represent their farness.

5. Conclusion

We employed Facet-SS-3D-Clump to identify and verify molecular clumps in the ^{13}CO data of MWISP, covering the longitude range $10° \leqslant l \leqslant 20°$ and a latitude strip of $|b| \leqslant 5°.25$. Through the analysis of the results, we conclude that: (1) We obtained a catalog containing 18,575 ^{13}CO clumps from the MWISP data using Facet-SS-3D-Clump. (2) The flux completeness limits showed that the catalog is 90% complete above 37 K km s^{-1}. (3) There are 82.3% of the dust clumps corresponding to ^{13}CO clumps from matching results with the Herschel infrared dust clumps. The Facet-SS-3D-Clump will be applied to survey ^{13}CO clumps across the entire MWISP data.

Acknowledgments

Appendix A
Examples of Clumps verified by SS-3D-Clump

To provide a demonstration of the efficacy of SS-3D-Clump in verifying molecular clump candidates, a subset of the samples were visualized. Figure A1 shows molecular clumps on the *l*–*b* plane with confidence levels obtained through SS-3D-Clump verification.
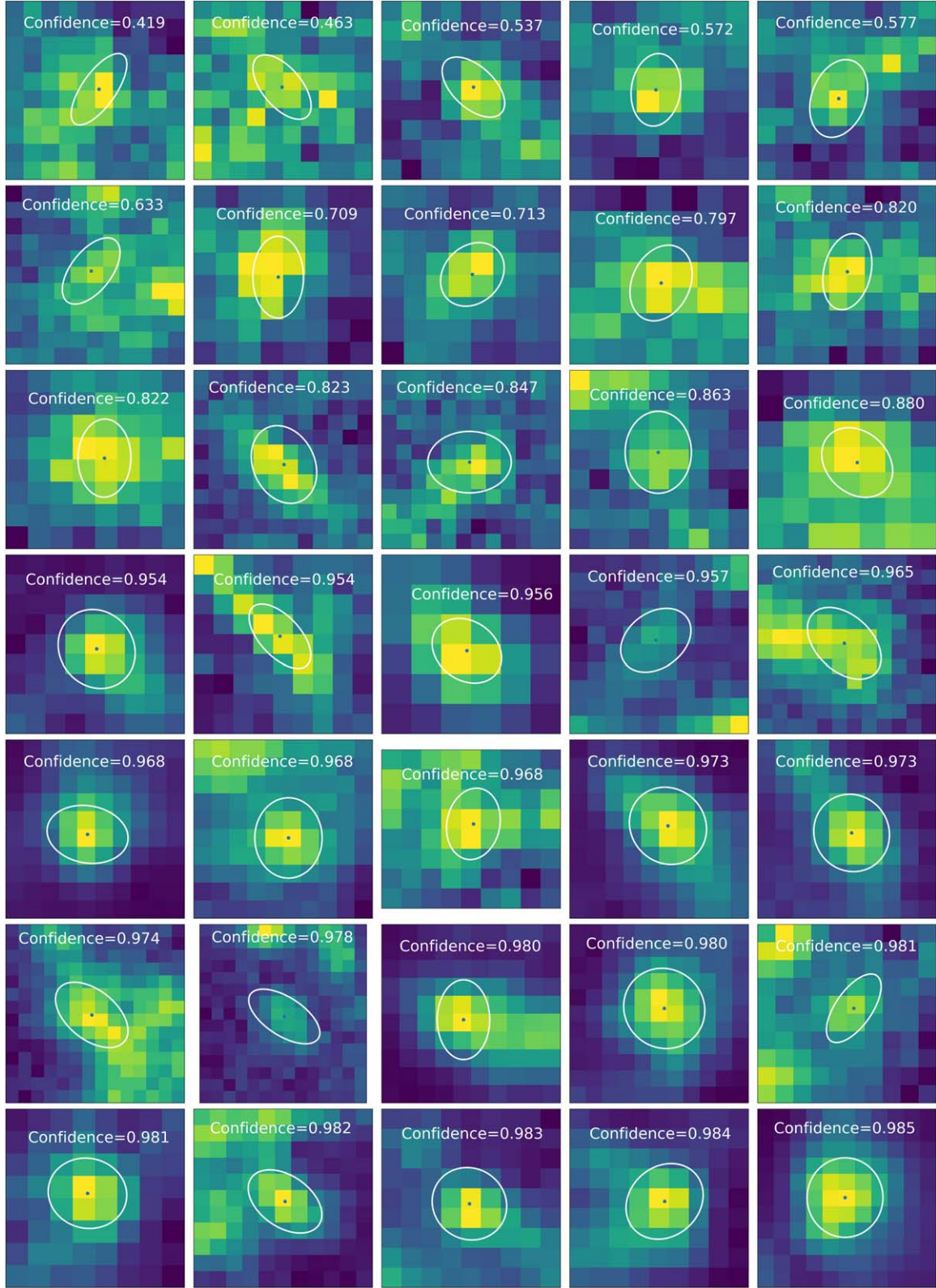
Figure A1. The integrated maps on the $l - b$ plane of the molecular clumps are presented. In the figures, white numbers indicate the confidence assigned by SS-3D-Clump for the current clump, blue dots represent the centroids of the clumps, and white ellipses depict the results of the 3D Gaussian modeling performed by MGM for the clump. The white ellipses mark the positions of the Gaussian ellipsoidal FWHM.

Appendix B
Partial Molecular Clumps in the Catalog

Table B1 presents a partial catalog of the clumps, providing information about their positions, sizes, intensities, and fluxes.

Table B1
Partial Catalog of Molecular Clumps

ID	l	b	v_{LSR}	S_M	S_m	S_v	θ	T_{peak}	W	E	α
MWSIP012.227-00.031+032.3	12.227	−0.031	32.3	136	101	1.18	7	3	65	1.32	0.99
MWSIP012.091-00.069+013.1	12.091	−0.069	13.1	148	101	1	180	2.2	42.7	0.69	0.99
MWSIP011.845-00.100+015.6	11.845	−0.1	15.62	160	118	1.63	180	6.4	261.8	1.34	0.99
MWSIP015.583+00.269+035.5	15.583	0.269	35.5	142	93	0.43	38	1.1	8.5	0.48	0.57
MWSIP012.505-00.236+004.3	12.505	−0.236	4.31	168	124	0.85	90	1.2	29.6	0.57	0.85
MWSIP012.033-00.121+022.2	12.033	−0.121	22.16	211	99	0.89	27	0.9	21.3	0.95	0.96
MWSIP012.200-00.039+050.5	12.2	−0.039	50.47	135	124	1.9	136	6.4	268.8	1.22	0.97
MWSIP011.879-00.087+050.3	11.879	−0.087	50.27	121	88	1.32	125	1.4	25.6	0.59	0.94
MWSIP015.467+00.272+029.0	15.467	0.272	28.98	124	100	0.98	180	1	16.4	0.53	0.86
MWSIP012.010-00.181+038.9	12.01	−0.181	38.94	262	176	1.43	132	3.8	338	1.46	0.99
MWSIP012.031-00.032+110.7	12.031	−0.032	110.71	125	105	0.87	48	6.4	96.7	1.32	0.98
MWSIP012.031-00.250+117.4	12.031	−0.25	117.39	113	103	1.37	118	2.1	45.5	0.70	0.99
MWSIP011.776-00.165+044.4	11.776	−0.165	44.4	134	61	0.69	35	1.8	13.5	1.12	0.98
MWSIP011.848+00.005+012.3	11.848	0.005	12.32	198	116	1.09	129	1.6	52.5	0.74	0.97
MWSIP012.093-00.102+113.1	12.093	−0.102	113.09	188	104	0.98	129	1.7	42.7	0.84	0.99
MWSIP017.513-01.114+043.3	17.513	−1.114	43.3	130	66	0.33	140	1.0	3.6	0.57	0.46
MWSIP011.966+00.050+018.9	11.966	0.05	18.91	151	93	1.16	37	2.6	55.5	0.77	0.99
MWSIP012.058-00.142+019.6	12.058	−0.142	19.58	135	91	0.71	20	1.2	14	0.54	0.58
MWSIP012.060-00.150+000.8	12.06	−0.15	0.77	135	113	0.76	180	2.1	32.4	0.79	0.97
MWSIP012.133+00.074+017.8	12.133	0.074	17.81	134	128	1.11	90	2.4	60.3	0.97	0.99
MWSIP011.766-00.068+060.4	11.766	−0.068	60.38	134	88	0.77	180	7.1	86.1	1.68	0.98
MWSIP011.864-00.231+046.4	11.864	−0.231	46.39	129	93	0.78	54	1.3	15.6	0.63	0.96
MWSIP012.190+00.036+048.5	12.19	0.036	48.49	174	128	1.26	18	3.5	131.9	1.13	1.00
MWSIP012.047+00.006+061.5	12.047	0.006	61.47	75	59	0.72	126	3.2	13.2	0.70	0.96
MWSIP012.129+00.059+112.4	12.129	0.059	112.38	119	109	1.27	90	1.6	36.2	0.64	0.97
MWSIP012.213+00.014+031.4	12.213	0.014	31.43	150	136	1.12	7	3.5	105.6	0.95	0.99
MWSIP012.174+00.227+029.2	12.174	0.227	29.18	109	80	0.83	11	1.3	12.9	0.76	0.97
MWSIP013.909-04.638+011.4	13.909	−4.638	11.36	105	71	0.37	1	1.6	5.9	0.55	0.82
MWSIP012.016-00.210-003.5	12.016	−0.21	-3.48	137	96	1.79	100	6.5	203	1.05	0.97
MWSIP018.812+00.272+018.5	18.812	0.272	18.5	140	90	1.03	90	3	51.8	0.99	0.99
MWSIP013.731-00.799+019.3	13.731	−0.799	19.28	156	111	0.91	120	1.8	37.4	0.65	0.99
MWSIP012.201-00.117+027.1	12.201	−0.117	27.06	147	140	2.14	1	8.9	521.5	1.47	0.98
MWSIP011.884-00.131+018.8	11.884	−0.131	18.77	115	107	1.39	124	1.8	42.1	0.80	0.98
MWSIP011.964-00.136+032.6	11.964	−0.136	32.56	153	145	0.94	130	4	112.3	0.80	0.98
MWSIP014.313-00.825+013.4	14.313	−0.825	13.44	137	87	0.88	32	2.1	28.7	0.79	0.99
MWSIP018.358-00.006+125.3	18.358	−0.006	125.33	113	93	0.61	113	1.5	13	0.55	0.87
MWSIP012.231+00.193+015.6	12.231	0.193	15.62	187	162	1.29	106	2.9	152.2	0.89	1.00
MWSIP011.970-00.209+120.2	11.97	−0.209	120.15	110	86	0.96	33	2.1	25.3	0.64	0.98
MWSIP012.118+00.033+095.7	12.118	0.033	95.69	188	132	1.14	151	2.1	79.9	0.72	0.99
MWSIP011.835-00.191+001.6	11.835	−0.191	1.65	150	71	0.84	45	1.6	18.5	0.66	0.99
MWSIP011.904-00.251-003.9	11.904	−0.251	-3.91	133	98	1.39	1	2	46.9	0.68	0.99
MWSIP016.647-01.261+016.8	16.647	−1.261	16.76	148	116	0.86	109	3.4	66.4	0.86	0.99
MWSIP012.106-00.027+111.6	12.106	−0.027	111.58	138	106	1.25	124	3.7	89.5	1.01	1.00
MWSIP011.967+00.119+097.8	11.967	0.119	97.8	180	98	0.81	43	2.1	39.4	0.67	0.99
MWSIP011.789-00.013+037.4	11.789	−0.013	37.35	174	112	1.02	133	1.3	33.2	1.34	0.97
MWSIP014.400-00.824+021.4	14.4	−0.824	21.44	141	115	1.2	180	7	181.8	1.93	0.99
MWSIP014.277-00.944+021.8	14.277	−0.944	21.77	132	125	0.79	148	5.7	98.8	0.95	0.97
MWSIP011.870-00.110+004.1	11.87	−0.11	4.07	161	124	1.26	90	2.2	71.9	0.70	1.00
MWSIP012.194-00.024+014.1	12.194	−0.024	14.12	153	94	1.14	114	2.2	49.5	0.75	1.00
MWSIP015.287+00.257+028.1	15.287	0.257	28.15	163	124	0.56	155	1.4	21.4	0.49	0.82

Table B1
(Continued)

ID	l	b	v_{LSR}	S_M	S_m	S_v	θ	T_{peak}	W	E	α
MWSIP010.072+01.092+027.7	10.072	1.092	27.69	111	89	0.97	142	1.3	17.1	0.51	0.84
MWSIP019.224+00.385+020.2	19.224	0.385	20.22	152	95	0.94	125	3.6	65.6	1.45	0.98
MWSIP012.150-00.024+047.0	12.15	−0.024	46.97	168	126	1.09	154	3.9	120.4	1.20	1.00
MWSIP011.920+00.068+095.8	11.92	0.068	95.77	181	128	1.11	180	2	69.9	0.66	0.99
MWSIP011.915+00.012+034.2	11.915	0.012	34.25	192	124	1.14	147	3	106.5	1.09	0.99
MWSIP012.123+00.226+049.3	12.123	0.226	49.33	124	116	1.41	180	2.2	60.5	0.64	0.99
MWSIP016.324-00.218+025.7	16.324	−0.218	25.67	115	63	0.62	133	1.7	10.4	0.52	0.54
MWSIP014.495-00.873+016.0	14.495	−0.873	16.02	204	97	0.96	123	1.8	46.7	0.71	0.98
MWSIP012.253-00.219+047.5	12.253	−0.219	47.49	104	97	0.72	90	2.3	22.2	0.60	0.97
MWSIP016.573+00.042+045.3	16.573	0.042	45.27	133	67	0.65	156	1.4	10.7	0.53	0.84
MWSIP017.647-00.796+057.6	17.647	−0.796	57.63	98	66	0.66	8	1.5	8.6	0.57	0.57
MWSIP012.231+00.105+121.0	12.231	0.105	121.05	110	68	0.77	0	1.2	9.2	0.55	0.71
MWSIP012.207+00.169+049.9	12.207	0.169	49.91	106	94	0.83	103	2.3	26.1	0.65	0.98
MWSIP015.841+00.936+028.3	15.841	0.936	28.32	93	68	0.62	25	1.3	6.9	0.55	0.80
MWSIP013.732-00.896+021.0	13.732	−0.896	21.01	178	142	1.3	16	3.9	169.7	0.82	0.99
MWSIP010.779-00.139+020.1	10.779	−0.139	20.11	130	79	0.8	13	1.1	12.2	0.54	0.82
MWSIP012.241-00.148+014.4	12.241	−0.148	14.36	124	111	0.82	141	7	106	0.91	0.95
MWSIP018.287+00.129+033.6	18.287	0.129	33.62	94	76	0.73	45	1.6	10.7	0.51	0.71
MWSIP011.798+00.076+015.8	11.798	0.076	15.82	147	108	0.94	35	4	80.4	0.98	0.99

ORCID iDs

Xiaoyu Luo ⬤ https://orcid.org/0000-0003-0592-3042
Zhibo Jiang ⬤ https://orcid.org/0000-0002-5920-031X
Zhiwei Chen ⬤ https://orcid.org/0000-0003-0849-0692

References

Abadi, M., Agarwal, A., Barham, P., et al., 2015 v2.14.0 TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, Zenodo,
Alex Rodriguez, A. L. 2014, Science, 344, 1492
Alves, J., Lombardi, M., & Lada, C. J. 2007, A&A, 462, L17
André, P., Di Francesco, J., Ward-Thompson, D., et al. 2014, in Protostars and Planets VI, ed. H. Beuther et al. (Tucson, AZ: Univ. of Arizona Press), 27, Protostars and Planets VI
Arzoumanian, D., André, P., Könyves, V., et al. 2019, A&A, 621, A42
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Barnes, P. J., Muller, E., Indermuehle, B., et al. 2015, ApJ, 812, 6
Basu, S. 2016, in IAU Symposium, Vol. 315, From Interstellar Clouds to Star-Forming Galaxies: Universal Processes?, ed. P. Jablonka, P. André, & F. van der Tak (Cambridge: Cambridge University Press), 85
Basu, S., Banerjee, A., & Mooney, R. J. 2002, Machine Learning, Proc. of the Nineteenth Int. Conf. (ICML 2002), Univ. of New South Wales, Sydney, Australia, July 8-12, 2002
Benedettini, M., Molinari, S., Baldeschi, A., et al. 2017, MMSAI, 88, 722
Benedettini, M., Molinari, S., Baldeschi, A., et al. 2020, A&A, 633, A147
Benedettini, M., Pezzuto, S., Schisano, E., et al. 2018, A&A, 619, A52
Benedettini, M., Traficante, A., Olmi, L., et al. 2021, A&A, 654, A144
Bergin, E. A., & Tafalla, M. 2007, ARA&A, 45, 339
Berry, D. S. 2015, A&C, 10, 22
Beuther, H., Linz, H., Henning, T., et al. 2011, A&A, 531, A26
Blitz, L., & Stark, A. A. 1986, ApJL, 300, L89
Bresnahan, D., Ward-Thompson, D., Kirk, J. M., et al. 2018, A&A, 615, A125
Chabrier, G. 2003, PASP, 115, 763

Chen, Z., Sun, W., Chini, R., et al. 2021, ApJ, 922, 90
Cheng, T.-Y., Conselice, C. J., Aragón-Salamanca, A., et al. 2020, MNRAS, 493, 4209
Cheng, Y., Tan, J. C., Liu, M., et al. 2018, ApJ, 853, 160
Comrie, A., Wang, K.-S., Hsu, S.-C., et al., 2021 v2.0.0 CARTA: The Cube Analysis and Rendering Tool for Astronomy, Zenodo,
Du, X., Xu, Y., Yang, J., & Sun, Y. 2017, ApJS, 229, 24
Elia, D., Molinari, S., Schisano, E., et al. 2017, MNRAS, 471, 100
Felli, M., Churchwell, E., & Massi, M. 1984, A&A, 136, 53
Gama, D. R. G., Lepine, J. R. D., Mendoza, E., Wu, Y., & Yuan, J. 2016, ApJ, 830, 57
Gupta, R., Srijith, P., & Desai, S. 2022, A&C, 38, 100543
Hacar, A., Tafalla, M., Kauffmann, J., & Kovács, A. 2013, A&A, 554, A55
He, Z., Qiu, B., Luo, A.-L., et al. 2021, MNRAS, 508, 2039
Hill, T., Motte, F., Didelon, P., et al. 2012, A&A, 542, A114
Jackson, J. M., Rathborne, J. M., Shah, R. Y., et al. 2006, ApJS, 163, 145
Jiang, Y., Chen, Z., Zheng, S., et al. 2023, ApJS, 267, 32
Jiang, Z., & Li, J. 2013, Protostars and Planets VI Posters
Kerton, C. R., Arvidsson, K., & Alexander, M. J. 2013, AJ, 145, 78
Könyves, V., André, P., Men'shchikov, A., et al. 2010, A&A, 518, L106
Könyves, V., André, P., Men'shchikov, A., et al. 2015, A&A, 584, A91
Kroupa, P., Weidner, C., Pflamm-Altenburg, J., et al. 2013, in Planets, Stars and Stellar Systems. Volume 5: Galactic Structure and Stellar Populations, ed. T. D. Oswalt & G. Gilmore, 5 (Dordrecht: Springer), 115
Krumholz, M. R., & McKee, C. F. 2005, ApJ, 630, 250
Krumholz, M. R., McKee, C. F., & Tumlinson, J. 2009, ApJ, 699, 850
Lada, E. A. 1992, ApJL, 393, L25
Liu, L., Bureau, M., Li, G.-X., et al. 2022, MNRAS, 517, 632
Lukic, V., de Gasperin, F., & Brüggen, M. 2019, Galax, 8, 8, Galaxies
Luo, X., Zheng, S., Huang, Y., et al. 2022, RAA, 22, 015003
Luo, X., Zheng, S., Jiang, Z., et al. 2024, A&A, 683, A104
Maity, A. K., Dewangan, L. K., Sano, H., et al. 2022, ApJ, 934, 2
Marsh, K. A., Kirk, J. M., André, P., et al. 2016, MNRAS, 459, 342
Messineo, M., Clark, J. S., Figer, D. F., et al. 2015, ApJ, 805, 110
Molinari, S., Schisano, E., Elia, D., et al. 2016, A&A, 591, A149
Molinari, S., Swinyard, B., Bally, J., et al. 2010, PASP, 122, 314
Motte, F., Andre, P., & Neri, R. 1998, A&A, 336, 150

Motte, F., Bontemps, S., & Louvet, F. 2018, ARA&A, 56, 41
Mottram, J. C., & Brunt, C. M. 2010, in ASP Conf. Ser. 438, The Dynamic Interstellar Medium: A Celebration of the Canadian Galactic Plane Survey, ed. R. Kothes, T. L. Landecker, & A. G. Willis (San Francisco, CA: ASP), 98
Nakanishi, H., Fujita, S., Tachihara, K., et al. 2020, PASJ, 72, 43
Nejad-Asghar, M. 2011, Ap&SS, 334, 27
Nutter, D., & Ward-Thompson, D. 2007, MNRAS, 374, 1413
Ohashi, S., Sanhueza, P., Chen, H.-R. V., et al. 2016, ApJ, 833, 209
Olmi, L., Ade, P. A. R., Anglés-Alcázar, D., et al. 2009, ApJ, 707, 1836
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, JMLR, 12, 2825
Pety, J. 2005, SF2A-2005: Semaine de l'Astrophysique Francaise ed. F. Casoli et al. (Paris: EdP Sciences), 721
Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1
Rani, R., Moore, T. J. T., Eden, D. J., et al. 2023, MNRAS, 523, 1832
Rathborne, J. M., Lada, C. J., Muench, A. A., et al. 2009, ApJ, 699, 742
Rigby, A. J., Moore, T. J. T., Eden, D. J., et al. 2019, A&A, 632, A58
Rigby, A. J., Moore, T. J. T., Plume, R., et al. 2016, MNRAS, 456, 2885
Rojas, K., Savary, E., Clément, B., et al. 2022, A&A, 668, A73
Rosolowsky, E. W., Pineda, J. E., Kauffmann, J., & Goodman, A. A. 2008, ApJ, 679, 1338
Schuller, F., Csengeri, T., Urquhart, J. S., et al. 2017, A&A, 601, A124
Shan, W., Yang, J., Shi, S., et al. 2012, ITTST, 2, 593
Shu, F. H., Adams, F. C., & Lizano, S. 1987, ARA&A, 25, 23
Su, Y., Yang, J., Zhang, S., et al. 2019, ApJS, 240, 9
Sun, J. X., Lu, D. R., Yang, J., et al. 2018, AcASn, 59, 3
Sun, Y., Su, Y., Zhang, S.-B., et al. 2017, ApJS, 230, 17
Sun, Y., Xu, Y., Yang, J., et al. 2015, ApJL, 798, L27
Takekoshi, T., Fujita, S., Nishimura, A., et al. 2019, ApJ, 883, 156
Tremblin, P., Schneider, N., Minier, V., et al. 2014, A&A, 564, A106
Tursun, K., Henkel, C., Esimbek, J., et al. 2022, A&A, 658, A34
Williams, J. P., Blitz, L., & McKee, C. F. 2000, in Protostars and Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell (Cambridge: Cambridge University Press), 97
Wu, Y., Liu, T., Meng, F., et al. 2012, ApJ, 756, 76
Wurster, J., & Rowan, C. 2023, MNRAS, 523, 3025
Yin, J., Chen, Z., Yao, Y., et al. 2022, RAA, 22, 035021
Zhang, Q., Wang, Y., Pillai, T., & Rathborne, J. 2009, ApJ, 696, 268
Zhang, S., Xu, Y., & Yang, J. 2014, AJ, 147, 46
Zhang, S., Zavagno, A., López-Sepulcre, A., et al. 2021, A&A, 646, A25
Zhu, X.-P., Dai, J.-M., Bian, C.-J., et al. 2019, Ap&SS, 364, 55
Zinnecker, H., & Yorke, H. W. 2007, ARA&A, 45, 481
Zuo, Y.-X., Li, Y., Sun, J.-X., et al. 2011, ChA&A, 35, 439